

MAKING EVALUATIONS RELEVANT TO
CONGRESSIONAL NEEDS

INVITED ADDRESS TO THE AMERICAN PSYCHOLOGICAL ASSOCIATION'S
89TH ANNUAL CONVENTION, LOS ANGELES, AUGUST 24-28, 1981

ELEANOR CHELIMSKY
DIRECTOR, INSTITUTE FOR
PROGRAM EVALUATION
THE GENERAL ACCOUNTING OFFICE

PRESIDENT, THE EVALUATION
RESEARCH SOCIETY

IT'S A PLEASURE FOR ME TO BE HERE TODAY, SPEAKING TO THIS GROUP WHICH HAS DONE SO MUCH IN THE PAST TO ADVANCE THE QUALITY, THE SPREAD, AND THE PRESTIGE OF EVALUATION. I'D LIKE TO SHARE WITH YOU THIS MORNING SOME NOTIONS I HAVE ABOUT MAKING EVALUATIONS USEFUL, AND ABOUT MAKING THEM USEFUL IN A PARTICULAR ARENA, THAT OF LEGISLATIVE POLICYMAKING.

OF COURSE, THE PROBLEM OF EVALUATION USE IS HARDLY A NEW ONE. WE'VE BEEN PUZZLING FOR YEARS NOW ABOUT HOW TO GET PROGRAM MANAGERS, FOR EXAMPLE, TO USE EVALUATION FINDINGS AS A WAY OF CORRECTING OR EVEN CUTTING THEIR PROGRAMS, DESPITE THE OBVIOUS CAREER THREAT WHICH SUCH USE CAN REPRESENT. OR ABOUT HOW TO GET HARRIED ADMINISTRATORS (WHO DISPOSE OF MAYBE 15 MINUTES AT THE END OF A DAY) TO READ EVALUATIONS, DESPITE THEIR TYPICAL LENGTHINESS, COMPLEXITY, AND SEEMINGLY INEVITABLE INFESTATIONS OF JARGON. OR ABOUT HOW TO GET PRACTITIONERS TO MODIFY TRADITIONAL PRACTICE BASED ON EVALUATION FINDINGS, DESPITE THE OFTEN HIGHLY UNSPECIFIC AND DUBIOUSLY APPLICABLE NATURE OF THE RECOMMENDATIONS. OR ABOUT HOW TO GET LEGISLATORS TO CONSIDER EVALUATION FINDINGS IN, SAY, PROGRAM REAUTHORIZATIONS, DESPITE THE SOMETIMES DISTANT RELEVANCE OF THE FINDINGS TO THE CONTROVERSY AT HAND, OR THE FAILURE OF THE EVALUATION PRODUCTS TO ARRIVE ON TIME.

NOW, AS WE EVALUATORS HAVE BEEN HEARING MORE AND MORE ABOUT THE DIFFICULTIES WHICH USERS HAVE HAD IN APPLYING EVALUATION FINDINGS, WE'VE CONCENTRATED MORE AND MORE ON THE EVALUATIONS THEMSELVES, TRYING TO MAKE THEM MORE RIGOROUS, MORE PERFECT, ON THE ASSUMPTION THAT IT WAS PERHAPS POOR QUALITY WHICH IMPEDED USE. BUT THE FACT IS, AS WE ALL KNOW, THAT SOME VERY MEDIOCRE EVALUATIONS HAVE BEEN

USED VERY WELL INDEED, WHILE SOME VERY EXCELLENT WORK HAS LANGUISHED ELEGANTLY ON THE SHELVES AND IN-BASKETS OF POLICY-MAKERS, -- WHETHER OR NOT THEY SERVED TO HIGHLIGHT METHODOLOGICAL POINTS AMONG RESEARCH AUDIENCES.

AS EVALUATORS, IN SUM, UNDER CRITICISM, WE'VE CONTINUED TO SCRUTINIZE THE SUBJECT OF OUR INTEREST -- THAT IS, THE EVALUATION -- RATHER THAN THE NEEDS OF THE USER.

NOW, OF COURSE, THERE'S NOTHING VERY SURPRISING ABOUT THAT, GIVEN THAT EVALUATION RESEARCH HAS MANY OF THE CHARACTERISTICS OF RESEARCH IN GENERAL. THAT IS, THE RESEARCHER, AS SUCH, IS INTERESTED IN PRODUCING NEW KNOWLEDGE, KNOWLEDGE FOR ITS OWN SAKE. AS WE KNOW, RESEARCHERS ARE TRAINED WITHIN PARTICULAR DISCIPLINES AND THEIR WORK TYPICALLY FLOWS FROM IDEAS OR HYPOTHESES WHICH THEY DEVELOP, AND WHICH USUALLY BUILD UPON THE IDEAS AND HYPOTHESES OF RESEARCHERS BEFORE THEM. THEY TEST THESE HYPOTHESES AS SYSTEMATICALLY AS THEY CAN, AND THEY CAREFULLY DOCUMENT THOSE TESTS AND THE FINDINGS WHICH ENSUE. THEIR EXPECTATION IS THAT BOTH TESTS AND FINDINGS WILL EITHER BE BORNE OUT OR THROWN OUT BY OTHER RESEARCHERS, WHO WILL THEN BUILD UPON THE NEW FINDINGS, USING THE SAME SYSTEMATIC, FORMAL METHODS. IT'S THIS FAMILIAR STEP-BY-STEP, CAREFULLY EXPLICATED, ITERATIVE BUT NEARLY CLOSED-CIRCUIT PROCESS -- IN WHICH THE RESEARCHER IS BOTH PRODUCER AND USER -- THAT CONFERS UPON THE RESEARCH ENTERPRISE THREE OF ITS ESSENTIAL CHARACTERISTICS: COOPERATION, CUMULATION, AND REVIEW (OR USE) BY PEERS.

A FOURTH CHARACTERISTIC IS CONTROL. THAT IS, RESEARCHERS EXPECT TO HAVE GREAT INDEPENDENCE IN DESIGNING AND EXECUTING THEIR WORK, AT LEAST PARTLY BECAUSE THEY NEED TO ENSURE CONTINUITY IN THE THINKING WHICH PRODUCES THE RESEARCH DESIGN AND GUIDES THE ANALYSIS AND INTERPRETATION OF RESULTS. SO ALMOST TOTAL CONTROL AND DIRECTION BY RESEARCHERS HAVE LONG BEEN THE NORM IN THE PERFORMANCE OF THE RESEARCH FUNCTION.

BUT THESE PARTICULAR CHARACTERISTICS -- "TRUTH FOR TRUTH'S SAKE," DEPENDENCE UPON EARLIER RESEARCH, FOCUS ON OTHER RESEARCHERS AS THE ESSENTIAL RESEARCH USER, AND AUTONOMY IN THE CONDUCT OF THE WORK -- HAVE A NATURAL, BUT PERHAPS UNINTENDED SIDE-EFFECT. BECAUSE THEY FOCUS SO HEAVILY UPON THE RESEARCH (OR UPON THE EVALUATION), AND BECAUSE EVERY EMPHASIS ALSO IMPLIES SOME DEPRIVATION, THESE CHARACTERISTICS LEAD INEVITABLY TO AT LEAST SOME INDIFFERENCE ON THE RESEARCHER'S PART IN THE USE OR APPLICATION OF RESEARCH FINDINGS WHEN THE USERS ARE NOT RESEARCHERS.

IN TRYING TO UNDERSTAND AND IMPROVE THE USE THAT'S MADE OF RESEARCH FINDINGS, THEN, IT'S IMPORTANT TO BEGIN WITH THE RESEARCHER'S PERSPECTIVE. IN BERANEK'S WORDS, ALTHOUGH THE MAJORITY OF RESEARCHERS MAY NEVER MAKE SIGNIFICANT CONTRIBUTIONS NOR ACHIEVE RECOGNITION, THEY STILL WORK AS RESEARCHERS BECAUSE OF WHAT THEY SEE AS THE NOBILITY, THE WORTHWHILENESS, THE "CHARISMA" OF RESEARCH. FOR MOST RESEARCHERS, THE VERY ACT OF DOING RESEARCH FULFILLS THEIR PUBLIC OBLIGATION. AND THE PEER

REVIEW SYSTEM REWARDS THEM FOR DOING GOOD RESEARCH, NOT FOR WHATEVER USE IS LATER MADE OF THEIR FINDINGS. ^{1/}

BUT THE FACT THAT THE "USE ISSUE" FAILS TO RECEIVE A GREAT DEAL OF ATTENTION FROM RESEARCHERS IS NOT THE ONLY PROBLEM. MANY SOCIAL RESEARCHERS, MANY EVALUATORS, ACTUALLY EQUATE A FOCUS ON USE WITH WHAT THEY CALL "CLIENT-SATISFICING," OR THE UNFORTUNATE AND IMPROPER SUBSTITUTION OF A MARKET GOAL FOR THE KNOWLEDGE GOAL WHICH IS BASIC TO THE ENTIRE RESEARCH ENTERPRISE. SO THE FACT THAT RESEARCHERS ARE INNER-DIRECTED, PEER-DIRECTED, AND KNOWLEDGE-DIRECTED -- CHARACTERISTICS THAT WE OFTEN HEAR REPRESENTED AS THE GREAT STRENGTH OF BOTH A PARTICULAR PIECE OF WORK AND OF THE "INTER-CONNECTED FABRIC" ^{2/} OF RESEARCH, IN HARVEY BROOKS' TERM -- THAT FACT CAN POSE PROBLEMS FOR POLICYMAKERS AND OTHERS NEEDING TO USE THE FINDINGS. FOR NOT ONLY DO RESEARCHERS TEND TO THINK FIRST OF THEIR OWN RESEARCH INTERESTS (AND THOSE OF THEIR RESEARCH FIELD), AND ONLY SECOND -- IF AT ALL -- OF THE NEEDS OF RESEARCH USERS, BUT FURTHER, AN EFFORT TO EMPHASIZE THE NEEDS OF USERS IS ACTUALLY CONSIDERED SOMEHOW UNWORTHY AND EVEN DANGEROUS BY MORE THAN A FEW RESEACHERS AND EVALUATORS.

WHAT DOES THIS MEAN IF WE'RE TRYING TO IMPROVE THE USE OF RESEARCH OR EVALUATION FINDINGS BY NON-RESEARCH AUDIENCES? WELL, IN THE FIRST PLACE, IT MEANS THAT RESEARCH AND NON-RESEARCH

^{1/}WILLIAM BERANEK, JR., CHOOSING AND USING SCIENTIFIC ADVICE IN THE STATE LEGISLATURE, HOLCOMB RESEARCH INSTITUTE, JUNE 1979.

^{2/}HARVEY BROOKS, "CAN SCIENCE SURVIVE IN THE MODERN AGE?" SCIENCE VOLUME 174, OCTOBER 1, 1972.

PERSPECTIVES (PERSPECTIVES OF POLICYMAKERS OR PRACTITIONERS, FOR EXAMPLE) ARE NOT NATURALLY LINKED, BUT NEED TO BE RECONCILED, IF FINDINGS ARE TO BE BOTH USABLE AND USED. SECOND, IT MEANS THAT THIS MAY NOT BE EASY TO DO BECAUSE THE COMBINATION OF RESEARCHER INTERDEPENDENCE, RESEARCH AUTONOMY, AND THE PURSUIT OF GOALS REQUIRING HIGH LEVELS OF TECHNICAL COMPETENCE, TENDS TO ISOLATE RESEARCHERS AS A GROUP. SUCH ISOLATION, WHICH MAY REALLY BE A FIFTH CHARACTERISTIC OF THE RESEARCH FUNCTION, ALLOWS A DEEP GULF OF MISUNDERSTANDING TO WIDEN AND HARDEN BETWEEN THE WORLD OF RESEARCH AND THE OUTSIDE WORLD 3/.

NOW, THE EXISTENCE OF THIS GULF HAS TWO EFFECTS. FIRST, IT GIVES RISE TO RESEARCHER VIEWS OF THE OUTSIDE WORLD AS (QUOTE) "HOSTILE, INCAPABLE OF ENCOMPASSING OBJECTIVE RESEARCH, AND BENT ON SUBORDINATING IT TO DUBIOUS GOALS SUCH AS PROFIT-MAKING, OR THE CREATION OF ARTIFICIAL NEEDS." 4/ SECOND, IT REINFORCES USER-HELD IMAGES OF EVALUATORS AND RESEARCHERS WHO MAY BE VIEWED AS FANATIC, INFLEXIBLE OR ARROGANT AND THEIR EVALUATION OR RESEARCH PRODUCTS AS "OVER-THEOLOGIZED, SPOOKY, UNFATHOMABLE, TEDIOUS TO READ, AND COMPLICATED TO UNDERSTAND," IN CAREY'S PHRASE. 5/

THIS SECOND EFFECT, WHICH SPEAKS TO THE IMAGE OF RESEARCHERS HELD BY NON-RESEARCHERS, BRINGS US TO THE OTHER IMPORTANT PERSPECTIVE

3/"POLICY FOR THE SOCIAL SCIENCES: AN EXAMINERS' REPORT," STANLEY HOFFMAN, WASSILY LEONTIEV, AND HENRY TAJFEL, IN THE OECD OBSERVER, OCTOBER-NOVEMBER 1974, No. 72, p. 21.

4/IBID.

5/ELEANOR CHELIMSKY, AN ANALYSIS OF THE PROCEEDINGS OF A SYMPOSIUM ON THE USE OF EVALUATION BY FEDERAL AGENCIES, THE MITRE CORPORATION, M77-39, JULY 1977, p. 43, REMARKS OF WILLIAM D. CAREY.

WE NEED TO UNDERSTAND IF WE WANT TO IMPROVE THE USE OF RESEARCH OR EVALUATION FINDINGS, AND THAT'S THE PERSPECTIVE OF THE USER. IN THIS CASE, TODAY, WE'RE TALKING ABOUT A LEGISLATIVE USER. WHAT DO WE REALLY KNOW ABOUT THE LEGISLATIVE PERSPECTIVE, ESPECIALLY INsofar AS ITS INTERACTIONS WITH THE RESEARCH PERSPECTIVE ARE CONCERNED?

TO BEGIN WITH, IT WOULD APPEAR THAT, IN THE NORMAL COURSE OF EVENTS, LEGISLATORS AND RESEARCHERS ARE SEPARATED BY DIFFERENT GOALS, DIFFERENT STANDARDS OF EVIDENCE, DIFFERENT TOLERANCES OF UNCERTAINTY. RESEARCHERS, AS WE'VE SAID, HAVE A GOAL OF KNOWLEDGE OR TRUTH; LEGISLATORS, ON THE OTHER HAND, HAVE A GOAL OF POWER OR IMPACT, AS DON PRICE HAS OBSERVED. 6/ NOW, THE RESEARCHER'S KNOWLEDGE GOAL IS, OF COURSE, IMPLEMENTED VIA THE SCIENTIFIC METHOD AND BOTH GOAL AND METHOD LEAD RESEARCHERS, FIRST, TO BELIEVE IN RATIONAL SOLUTIONS TO DILEMMAS AND, SECOND, TO DEFINE UNCERTAINTY VERY CAREFULLY. BUT THE LEGISLATIVE POWER GOAL, WHICH IS IMPLEMENTED VIA THE ADVERSARY PROCESS (NOT THE SCIENTIFIC METHOD), INSTEAD LEADS LEGISLATORS TO BELIEVE THAT THE BEST SOLUTION IS THE ONE PRESENTED WITH THE MOST PERSUASIVE ARGUMENTS OR BACKED UP BY THE MOST IMPRESSIVE AUTHORITY. 7/ THAT IS, FOR THE LEGISLATOR, THE EVIDENCE IS INSTRUMENTAL TO THE NEGOTIATION OR TO THE DECISION, WHEREAS FOR THE RESEARCHER, THE EVIDENCE IS THE END IN ITSELF.

A FEW YEARS AGO, THE NATIONAL SCIENCE FOUNDATION CONDUCTED AN IMPORTANT EFFORT TARGETING THE TRANSFER OF EVALUATIVE SCIENTIFIC

6/DON K. PRICE, THE SCIENTIFIC ESTATES, HARVARD PRESS, 1965, CHAPTER 5.

7/BERANEK (1979).

INFORMATION BETWEEN RESEARCHERS AND LEGISLATORS IN 42 STATES. RESEARCHERS WORKED DIRECTLY FOR LEGISLATORS, AND A RESEARCHER 8/ WORKING IN THE STATE OF INDIANA DESCRIBED THE PROBLEM THIS WAY:

THE RESEARCHER SAYS HE'S CERTAIN ONLY WHEN HE'S ABSOLUTELY CERTAIN. TO HIM, MAKING NO STATEMENT IS BETTER THAN MAKING A STATEMENT BASED ON INCOMPLETE DATA. BUT IN POLICYMAKING, POSITIVE VALUE IS PLACED ON "MAKING A DECISION," REGARDLESS OF WHETHER OR NOT THERE IS SUFFICIENT OBJECTIVE EVIDENCE TO SUPPORT THE DECISION. WE SEE THIS CONFLICT VIVIDLY IN RECENT ATTEMPTS TO ESTABLISH PERMISSIBLE CHEMICAL CARCINOGEN LEVELS IN FACTORY EMISSIONS, IN OCCUPATIONAL EXPOSURES, AND IN FOOD SUBSTANCES TO INSURE REASONABLE SAFETY. BUT PROVING THAT A SUBSTANCE CAUSES CANCER IN HUMANS IS VIRTUALLY IMPOSSIBLE BY SCIENTIFIC CRITERIA OF PROOF. THE BEST THAT CAN BE DONE IS TO SHOW THAT UNDER CERTAIN SPECIAL CIRCUMSTANCES, THE SUBSTANCE IS ASSOCIATED WITH THE INDUCTION OF CANCER IN ANOTHER SPECIES OF ANIMAL. THE SCIENTIST CANNOT EXTRAPOLATE THESE ANIMAL OBSERVATIONS TO HUMANS WITH SCIENTIFIC CERTAINTY. AND YET THE POLICY-MAKER MUST MAKE A DECISION ABOUT HUMANS. IN ORDER TO MAKE A TIMELY DECISION BASED ON THE BEST EVIDENCE AVAILABLE, CERTAINTY TO A POLICYMAKER CAN BE WHAT A SCIENTIST MIGHT CONJECTURE AS EITHER "DAMN LIKELY" OR "ALMOST NEVER."

FURTHER, THESE DIFFERENCES IN GOALS AND METHODS BETWEEN LEGISLATORS AND RESEARCHERS NATURALLY BRING DIFFERENCES IN PRIORITIES AND STRATEGIES, DIFFERENCES WHICH SIGNIFY THAT RESEARCHERS MAY SEE

8/BERANEK (1979).

THE NEED FOR RESEARCH EVERYWHERE AND LEGISLATORS MAY NOT SEE IT ANYWHERE. A MINNESOTA LEGISLATOR INVOLVED IN THE SAME NSF STUDY HAD THIS TO SAY:

LEGISLATIVE PRIORITIES EXIST IN THE FOLLOWING ORDER WITH REGARD TO A PARTICULAR ISSUE: FIRST, LEGISLATORS CONSIDER THE EFFECT ON CONSTITUENTS (HOW DO THEY FEEL ABOUT IT?); SECOND, THEY CONSIDER LEGISLATIVE FEASIBILITY (IS THERE A CONSENSUS TO DO SOMETHING ABOUT IT?); AND ONLY IN LAST PLACE DO THEY CONSIDER SUBSTANTIVE INFORMATION (WHAT DO WE KNOW ABOUT IT?). THE LEGISLATOR THUS REVERSES THE PRIORITIES OF THE IDEAL "STATESMAN-POLICYMAKER" WHO PUTS SUBSTANCE FIRST AND CONSTITUENTS LAST. ALSO, SOMEWHERE WITHIN THE LEGISLATOR'S LAST PRIORITY LIES THE RESEARCHER'S TOP PRIORITY. AND THIS EXPLAINS QUITE A FEW THINGS. FOR EXAMPLE: WHY EMOTIONAL ISSUES (I.E., CONSTITUENT ISSUES) DOMINATE LEGISLATURES; WHY, IF YOU COME IN WITH SUBSTANCE BUT WITHOUT SHOWING HOW CONSTITUENTS WILL BE AFFECTED OR WHAT LEGISLATIVE STRATEGY IS POSSIBLE, YOU WON'T GET MUCH RESPONSE; WHY SERVICE PACKAGES NEED TO BE CONCENTRATED ON GLAMOROUS ISSUES, NOT TECHNICAL ONES; AND WHY TIMEFRAMES FOR LEGISLATIVE ACTION ARE GEARED TO TIME IN OFFICE, NOT TO THE AMOUNT OF TIME NEEDED TO SOLVE A PROBLEM. 9/

BUT WHAT DO SUCH LEGISLATIVE PRIORITIES SIGNIFY FOR AN EVALUATOR? WELL, IF LEGISLATORS MUST THINK FIRST OF THEIR CONSTITUENTS, SECOND OF LEGISLATIVE FEASIBILITY AND ONLY THIRD

9/ THE HONORABLE GORDON VOSS, MINNESOTA STATE LEGISLATURE, REMARKS AT A PANEL DISCUSSION, CONFERENCE ON "THE INTEGRATION AND USE OF RESEARCH WITHIN THE FEDERAL SYSTEM," HAWAII, JUNE 19, 1979.

OF SUBSTANCE, AND IF WE ASSUME THAT THIS YEAR'S "EMOTIONAL" PROGRAM ISSUES MIGHT INCLUDE, SAY, FRAUD OR WASTE IN ENTITLEMENT PROGRAMS, COST GROWTH IN GOVERNMENT, POOR MANAGEMENT, "USELESS" PROGRAMS, AND PROGRAM INEFFECTIVENESS OR INEFFICIENCY, IN THAT ORDER, WE CAN SEE THAT EVALUATORS HAVE BEEN SPENDING MOST OF THEIR TIME ON THE LEAST EMOTIONAL ISSUES AND ADDRESSING CHIEFLY THE BOTTOM-MOST LEGISLATIVE PRIORITIES.

BUT IN ADDITION, THIS LOW LEGISLATIVE PRIORITY ASSIGNED TO SUBSTANCE IS REINFORCED BY THE TYPICALLY LEGAL, RATHER THAN SCIENTIFIC, TRAINING OF LAWMAKERS. THIS RESULTS IN A LACK OF LEGISLATIVE FAMILIARITY WITH RESEARCH MODES AND PROCEDURES, WHICH IN ITSELF DEEPENS THE ISOLATION BETWEEN LEGISLATORS AND RESEARCHERS AND CONSTITUTES AN IMPEDIMENT FOR LEGISLATIVE BODIES IN THE USE OF RESEARCH.

RELATIONSHIPS BETWEEN RESEARCHERS AND LEGISLATORS, THEN, ARE CONDITIONED NOT ONLY BY DIFFERENT TRAINING, DIFFERENT GOALS, DIFFERENT METHODS, AND DIFFERENT PRIORITIES, BUT ALSO AND ESPECIALLY BY DIFFERENT AREAS OF IGNORANCE. AND THE RESEARCHER IS USUALLY AS IGNORANT OF THE POLICY CONTEXT IN WHICH A RESEARCH QUESTION IS POSED AS THE LEGISLATOR MAY BE OF RESEARCH METHODS. YET TO ANSWER THE QUESTION, "IS BENZENE DANGEROUS?" THE RESEARCHER MUST KNOW IN WHAT CONTEXT; SIMILARLY, TO USE EVALUATION FINDINGS PROPERLY, THE LEGISLATOR MUST KNOW WHAT CONFIDENCE CAN BE PLACED IN THOSE FINDINGS AND HOW THEY CAN BE APPLIED.

TO SUM UP, IT SEEMS THAT AN EFFORT TO IMPROVE THE USE OF EVALUATION OR RESEARCH BY LEGISLATORS MUST TRY TO ADDRESS TWO KINDS OF PROBLEMS:

- FIRST, THE INATTENTION BY RESEARCHERS OR EVALUATORS TO THE PUBLIC USE OF THEIR WORK GENERALLY, AND TO THE INFORMATION NEEDS AND PRIORITIES OF LEGISLATORS, IN PARTICULAR; AND SECOND,
- PROBLEMS OF COMMUNICATION BETWEEN RESEARCHERS AND LEGISLATORS BECAUSE OF THEIR ISOLATION FROM EACH OTHER AND BECAUSE OF THEIR DIFFERENT PERSPECTIVES.

AT GAO'S INSTITUTE FOR PROGRAM EVALUATION, WHICH I DIRECT, ONE AMONG OUR SEVERAL MISSIONS IS TO MOUNT PRECISELY SUCH AN EFFORT WITH REGARD TO EVALUATION. WE HAVE BEGUN THIS EFFORT FIRST, BY TRYING TO GET AS CLEAR AN IDEA AS POSSIBLE OF LEGISLATIVE INFORMATION NEEDS; AND SECOND, BY TRYING TO BUILD TWO TYPES OF LINKAGES: LINKAGES OF LOGIC WHICH CAN BIND THOSE INFORMATION NEEDS TO PARTICULAR EVALUATION STRATEGIES; AND LINKAGES OF COMMUNICATION WHICH CAN IMPROVE UNDERSTANDING AND DEVELOP WORKING RELATIONSHIPS BETWEEN EVALUATORS AND CONGRESSIONAL STAFF.

FIRST, THEN, WHAT ARE THE MOST COMMON LEGISLATIVE INFORMATION NEEDS THAT CAN BE SATISFIED THROUGH EVALUATIVE TECHNIQUES? TO ANSWER THIS QUESTION, WE USED GAO, CONGRESSIONAL AND NSF SOURCES, AND DID SOME INTERVIEWS OF OUR OWN AS WELL. THE SENSE WE GOT WAS THAT THERE ARE AT LEAST SIX IMPORTANT WAYS IN WHICH EVALUATORS CAN RESPOND USEFULLY TO THE NEEDS OF LEGISLATORS AND LEGISLATIVE STAFF:

FIRST, OF COURSE, BY ANSWERING EVALUATIVE QUESTIONS, BUT ALSO, BY GETTING THE INFORMATION TO THE LEGISLATIVE USER RAPIDLY ENOUGH TO FIT THE TIME CONSTRAINTS OF THE CONGRESSIONAL NEGOTIATION OR DECISION PROCESS;

SECOND, BY SIFTING THROUGH THE QUANTITIES OF EXISTING EVALUATIVE INFORMATION TO SYNTHESIZE, ANALYZE, AND PRESENT SUCCINCTLY THAT WHICH IS RELEVANT TO THE QUESTION AT HAND;

THIRD, BY REFINING, WHERE POSSIBLE, THE FORMULATION OF THAT QUESTION-AT-HAND SO AS TO BE SURE OF TWO THINGS: THE FEASIBILITY OF OBTAINING THE INFORMATION SOUGHT, AND THE APPROPRIATENESS OF THAT INFORMATION TO THE POLICY NEED;

FOURTH, BY IDENTIFYING MAJOR GAPS IN AVAILABLE INFORMATION FOR FUTURE ATTENTION BY EVALUATORS AT GAO OR ELSEWHERE;

FIFTH, BY EVALUATING PLANS FOR NEW PROGRAMS OR POLICIES, AND REVIEWING EVALUATION REPORTS FOR OLD ONES; AND

SIXTH, BY DIRECTLY HELPING CONGRESSIONAL STAFF TO DEVELOP THE EVALUATIVE QUESTIONS ABOUT PROGRAM EFFECTIVENESS AND EFFICIENCY THAT ARE NEEDED IN THE LEGISLATIVE OVERSIGHT OF EXECUTIVE AGENCIES.

ASSUMING THAT THESE ARE THE FIRST AND MOST IMPORTANT KINDS OF LEGISLATIVE INFORMATION NEEDS THAT WE AS EVALUATORS SHOULD TRY TO FILL, WE ARE NOW SETTING UP OUR STRATEGY FOR DOING SO BY BUILDING LINKAGES OF LOGIC AND OF COMMUNICATION WITH THE CONGRESSIONAL USERS OF OUR WORK. THESE LINKAGES REST UPON TWO HYPOTHESES: FIRST, THAT DESIGNING BACKWARD FROM THE INFORMATION NEEDED IS FEASIBLE AND THAT IT WILL IN FACT ENSURE THE RELEVANCE, TIMELINESS, AND USE OF THE WORK PERFORMED; AND SECOND, THAT THE APPROPRIATENESS OF AN EVALUATION DESIGN TO THE INFORMATION NEEDING TO BE PRODUCED IS MORE IMPORTANT THAN THE RIGOR OF THAT DESIGN.

WITH THOSE IDEAS AS A STARTING POINT, OUR EVALUATION STRATEGY FOR WORKING WITH THE CONGRESS NOW CONTAINS THE FOLLOWING FIVE ELEMENTS:

ELEMENT #1: NEGOTIATING THE QUESTION IN TERMS OF THE INFORMATION NEED. WE BEGIN EVERY JOB BY A SERIES OF MEETINGS WITH THE CONGRESSIONAL SPONSORS TO REFINE THEIR QUESTION SO THAT WE CAN BE SURE, FIRST THAT IT'S A RESEARCHABLE ONE; SECOND, THAT IT CAN BE ANSWERED GIVEN THE TIME AND RESOURCES AVAILABLE WITH THE APPROPRIATE DEGREE OF CONCLUSIVENESS; AND THIRD, THAT THE CONGRESSIONAL USERS KNOW EXACTLY WHAT THEY'LL BE GETTING AND AGREE THAT IT'S WHAT THEY NEED. WE WON'T, FOR EXAMPLE, BEGIN A 2-YEAR JOB IF THE INFORMATION NEED DOESN'T REQUIRE IT OR IF WE ONLY HAVE 6 MONTHS' TIME; OR ANY JOB IF THE EFFORT DOESN'T SEEM LIKELY TO PRODUCE THE INFORMATION SOUGHT, OR IF WE DON'T GET SPONSOR AGREEMENT THAT THE INFORMATION IS INDEED THAT WHICH HE NEEDS.

ELEMENT #2: LEARNING MORE ABOUT THE EVALUATIVE INFORMATION NEEDS OF LEGISLATIVE USERS. WE NEED TO DEEPEN OUR UNDERSTANDING OF THE FORCES ACTING TO GENERATE EVALUATIVE INFORMATION NEEDS IN THE CONGRESS IF WE WANT TO RESPOND TO THEM OPTIMALLY.

TO BEGIN WITH, ALL INFORMATION NEEDS ARE NOT, OF COURSE, EVALUATIVE. IN ADDITION, SOME WHICH ARE EVALUATIVE MAY REQUIRE WORK OF A SORT WHICH MAKES THEM MORE APPROPRIATE FOR CONGRESSIONAL AGENCIES OTHER THAN THE GAO. SOME OTHER NEEDS MAY REQUIRE AUDITS RATHER THAN EVALUATIONS. SECONDLY, WE NEED TO BE ABLE TO DISTINGUISH AMONG GENERAL EVALUATIVE INFORMATION NEEDS SO AS TO ORGANIZE OUR RESPONSES USEFULLY ACCORDING TO SOME CLASSIFICATION. FOR EXAMPLE, THERE ARE EVALUATIVE NEEDS WHICH RECUR AT FIXED INTERVALS; NEEDS WHICH CALL FOR KNOWLEDGE OF PROGRAM EFFECTIVENESS; NEEDS WHICH REQUIRE ONLY DESCRIPTIVE INFORMATION; NEEDS WHICH MAY BE SUSCEPTIBLE TO A BUILDING-BLOCK, MIXED-METHODS STRATEGY; NEEDS WHICH ARE ANSWERABLE WITH EXTANT DATA, AND SO ON.

ELEMENT #3: DEVELOPING AND TESTING A PANOPLY OF EVALUATIVE TOOLS FOR USE WITH THE CONGRESS. GIVEN THAT MUCH OF THE CONGRESSIONAL CRITICISM LEVELLED AT EVALUATION HAS CONCERNED ITS TIMELINESS, WE MUST CONFRONT THE ^{TENSION BETWEEN THE} /LEGISLATIVE NEED FOR RAPID PROVISION OF EVALUATIVE INFORMATION AND THE FACT THAT DESIGNING AND EXECUTING RIGOROUS EVALUATIONS TAKES A LONG TIME. IN ADDITION, WE MUST ALSO CONFRONT THE TENSION BETWEEN THE GAO NEED TO MANAGE ITS WORK IN THE LEAST COSTLY WAY POSSIBLE (EVERY QUESTION SIMPLY

CANNOT BE ADDRESSED WITH A \$500,000 SURVEY OR A MILLION-DOLLAR EVALUATION) AND THE FACT THAT QUALITY IS SOMETIMES VERY EXPENSIVE. WE ARE THEREFORE USING THE NEGOTIATION PROCESS (JUST DISCUSSED AS ELEMENT #1) TO SEPARATE OUT THOSE QUESTIONS WHICH CAN BE APPROPRIATELY ANSWERED BY METHODS OTHER THAN RIGOROUS EVALUATIONS, AND WE ARE DEVELOPING OR ADAPTING AND TESTING SUCH METHODS FOR USE WITH LEGISLATIVE SPONSORS. LET ME REPORT TODAY ON THREE SUCH METHODS IN WHICH I THINK WE ARE MAKING SOME PROGRESS.

I. THE EVALUATION SYNTHESIS. AN EVALUATION SYNTHESIS BRINGS TOGETHER EXISTING STUDIES AND USES THEM AS A DATA BASE FOR ANSWERING SPECIFIC CONGRESSIONAL QUESTIONS. IT ALLOWS US TO DETERMINE (A) WHAT IS ACTUALLY KNOWN ABOUT A PARTICULAR TOPIC, (B) THE LEVEL OF CONFIDENCE ONE MAY HAVE IN THE VARIOUS STUDIES USED IN THE DATA BASE, AND (C) THE GAPS IN EVALUATIVE INFORMATION WHICH REMAIN.

WE DO AN EVALUATION SYNTHESIS EITHER WHEN A CONGRESSIONAL COMMITTEE ASKS FOR A STUDY OF THE EFFECTIVENESS OF A PROGRAM TO BE COMPLETED IN 9 MONTHS AND WE ARE ABLE TO NEGOTIATE THAT REQUEST TO THE SYNTHESIS AND EVALUATION OF EXISTING STUDIES; OR ELSE WHEN WE ARE ASKED FOR AN ANALYTICAL SUMMARY OF THE STATE OF THE EVALUATIVE ART IN A PARTICULAR AREA.

OUR FIRST EVALUATION SYNTHESIS WILL BE PUBLISHED IN SEPTEMBER. IT HAS ADDRESSED, FOR THE HOUSE SUBCOMMITTEE ON SELECT EDUCATION, THE TOPIC OF ACCESS TO SPECIAL EDUCATION BY THE HANDICAPPED. THE QUESTIONS WE AGREED TO ANSWER FOR THE SUBCOMMITTEE USING THE NEW EVALUATION SYNTHESIS METHOD WERE FOUR:

- WHAT ARE THE NUMBERS AND CHARACTERISTICS -- SUCH AS AGE, RACE, HANDICAPPING CONDITION AND ITS SEVERITY -- OF CHILDREN RECEIVING SPECIAL EDUCATION?
- WHAT ARE THE CHARACTERISTICS OF CHILDREN WHO ARE LESS OFTEN INCLUDED IN THE SPECIAL EDUCATION PROGRAM?
- WHAT ARE THE CHARACTERISTICS OF CHILDREN WHO MAY BE OVERREPRESENTED IN THE SPECIAL EDUCATION PROGRAM? AND
- WHAT ARE THE MAJOR FACTORS RELATED TO WHO GETS SPECIAL EDUCATION?

THE EVALUATION SYNTHESIS IS DESIGNED TO BE PERFORMED BY ONE OR TWO PERSONS, HAVING ACCESS TO SPECIALIZED EXPERTISE, OVER A TIME-PERIOD OF 6 TO 9 MONTHS. (THE PRECISE AMOUNT OF TIME WILL, OF COURSE, DEPEND ON THE NARROWNESS OF THE TOPIC AREA, AND THE SIZE OF THE DATA BASE AVAILABLE.)

WE DEVELOPED THIS METHOD TO RESPOND TO THREE NEEDS: THE NEED FOR RAPID EVALUATIVE INFORMATION, THE NEED FOR DIRECTING MASSIVE AMOUNTS OF EXISTING INFORMATION TO THE ANSWERING OF SPECIFIC QUESTIONS, AND THE NEED FOR A COST/EFFECTIVE USE OF IPE'S EVALUATIVE RESOURCES. IN ADDITION, THE METHOD ALLOWS IPE AT LEAST PARTIALLY TO FULFILL ITS MISSION OF IMPROVING THE USE WHICH THE CONGRESS MAKES OF EVALUATIVE INFORMATION SINCE, IN AND OF ITSELF, IT ENSURES THE INITIAL OR SECONDARY LEGISLATIVE USE OF EVALUATIONS WHICH HAVE ALREADY BEEN COMPLETED.

II. A SECOND METHOD WE ARE DEVELOPING FOR SERVING THE CONGRESS IS THE PODSE OR PROGRAM OPERATIONS AND DELIVERY OF SERVICES EXAMINATION. THIS IS AN ADAPTATION OF THE SERVICE DELIVERY ASSESSMENT DEVELOPED AT THE DEPARTMENT OF HEALTH AND HUMAN SERVICES. HERE WE ARE NOT SO FAR ADVANCED AS WITH THE EVALUATION SYNTHESIS, HAVING JUST COMPLETED A TENTATIVE METHODOLOGY WHICH WE'RE NOW GOING TO TEST. WITH THIS METHOD, WE WILL OURSELVES PERFORM A STUDY WHICH SEEKS TO ANSWER ONLY DESCRIPTIVE, NON-IMPACT TYPES OF QUESTIONS SUCH

AS: HOW ARE SERVICES BEING DELIVERED IN A PARTICULAR PROGRAM? WHAT DO BENEFICIARIES THINK OF THE PROGRAM? WHAT ARE THE PROGRAM'S VULNERABILITIES TO FRAUD AND ABUSE? IN DECIDING WHETHER A PODSE IS APPROPRIATE, WE LOOK AT WHAT INFORMATION ALREADY EXISTS AND AT THE KIND OF PROGRAM WHICH IS INVOLVED. A PODSE WOULD BE ESPECIALLY USEFUL, FOR EXAMPLE, IN A CASE WHERE A PROGRAM WAS SO NEW THAT LITTLE WAS KNOWN ABOUT IT, OR EXISTED ONLY TO SERVE PARTICULAR CLIENTS (SUCH AS CRIME VICTIM ASSISTANCE CENTERS OR VIETNAM VETERANS' OUTREACH PROGRAMS), SO THAT CLIENT AND PRACTITIONER OPINION OF THE PROGRAM WOULD BE CRUCIAL QUESTIONS, AND/OR WHERE SIGNIFICANT OPPORTUNITIES FOR SERVICE ABUSE MIGHT EXIST. A COUNTER-INDICATION FOR A PODSE, OF COURSE, WOULD BE THE NEED TO KNOW THE OVERALL EFFECTIVENESS OF A PROGRAM.

WE EXPECT THE PODSE TO INCLUDE SOME ELEMENTS OF AN OPINION SURVEY, COMBINED WITH SOME ASPECTS OF OPERATIONS ANALYSIS. IT WILL BE PERFORMED BY A TEAM OF 4 OR 5 PEOPLE WITH HELP FROM GAO'S REGIONAL OFFICES, TAKING WE THINK ABOUT 8-10 MONTHS TO COMPLETE.

WE ARE DEVELOPING THIS METHOD TO SERVE THE NEED FOR RAPID INFORMATION ON SERVICE DELIVERY BEING PERFORMED BY EXECUTIVE AGENCIES, AS WELL AS THE NEED FOR LEGISLATORS TO KNOW ABOUT AFFECTED CONSTITUENT ATTITUDES TOWARD SERVICE PROGRAMS. FINALLY, IT CAN EITHER STAND ALONE OR ESTABLISH THE NEED FOR A FULL-BORE EVALUATION OF PROGRAM EFFECTIVENESS.

III. THE THIRD METHOD I WANT TO MENTION IS THE EVALUATION PLANNING REVIEW. THIS METHOD IS STILL ON THE DRAWING BOARD, BUT IT SEEKS TO ANSWER LEGISLATIVE QUESTIONS ABOUT THE FEASIBILITY, THE LOGIC, AND THE REASONABLENESS OF PROPOSED NEW POLICIES AND PROGRAMS BASED ON THE EVALUATIVE EXPERIENCE OF THE PAST. THAT IS, BY GOING THROUGH THE EVALUATION PLANNING PROCESS FOR A NEW PROGRAM, LOOKING AT ITS OBJECTIVES AND THE RATIONALE GIVEN FOR BELIEVING THEY ARE REASONABLE, WE CAN IDENTIFY AREAS OF SOME UNCERTAINTY, OF GREAT UNCERTAINTY, AND OF TOTAL UNCERTAINTY. AT THE SAME TIME, WE CAN ENSURE THAT THOUGHT IS GIVEN TO THE LATER EVALUATION OF THE PROGRAM. THIS EFFORT ALLOWS CONGRESSIONAL STAFF TO ASK FOR MORE SUPPORTING DATA, IF THEY ARE NEEDED, OR TO TRY TO MODIFY THE OBJECTIVES, OR CHANGE THE IMPLEMENTATION OR ADMINISTRATION OF THE PROGRAM,

AND OF COURSE, IT SETS THE STAGE FOR THE EVALUATIVE QUESTIONS OF THE OVERSIGHT FUNCTION.

NOW THESE NEW METHODS NATURALLY DON'T EXCLUDE THE USUAL MIX WHICH IPE WILL PERFORM OF CASE STUDIES, SURVEYS, AND PROCESS OR OUTCOME EVALUATIONS. BUT AGAIN, WE WILL DO EXPENSIVE, LONG-TERM STUDIES ONLY WHEN THEY MUST BE DONE. THE BACKBONE OF THIS PROCESS IS THE UP-FRONT JOB NEGOTIATION WITH CONGRESSIONAL STAFF WHICH DETERMINES THE APPROPRIATE EVALUATION DESIGN FOR A PARTICULAR LEGISLATIVE INFORMATION NEED.

SO, WE HAVE THREE ELEMENTS OF A STRATEGY FOR DEVELOPING LINKAGES WITH THE CONGRESSIONAL USER OF EVALUATION FINDINGS: UP-FRONT NEGOTIATIONS, BETTER UNDERSTANDING BY US OF LEGISLATIVE EVALUATION NEEDS, AND NEW METHODS. THE FOURTH ELEMENT IS:

ELEMENT #4: FREQUENT BRIEFINGS OF CONGRESSIONAL STAFF AND CONTINUING COMMUNICATIONS ON PROGRESS. IT GOES WITHOUT SAYING TO ANYONE WHO HAS EVER CONDUCTED AN EVALUATION, THAT MANY THINGS CAN GO WRONG BETWEEN THE EVALUATOR'S TIDY JOB DESIGN AND THE MESSINESS OF JOB EXECUTION WHICH MAY OFTEN BE STUDDERED WITH PLEASANT EVENTS LIKE LOSING YOUR BEST STAFF, NOT FINDING THE DATA YOU COUNTED ON,

AND SO FORTH. BY BRIEFING THE LEGISLATIVE USER FREQUENTLY, WE HOPE TO ENSURE, FIRST, THAT THAT USER WILL KNOW WHAT WE'RE DOING, WHAT PROBLEMS WE'RE HAVING, AND WHAT CHANGES MAY BE NECESSARY IN THE FINAL PRODUCT WE HAD AGREED ON; AND SECOND, THAT WE WILL KNOW HOW THE USER'S INFORMATION NEED MAY HAVE CHANGED OR EXPANDED, ALLOWING US POSSIBLY TO BROADEN OUR OWN EFFORT, IF STUDY SCOPE, TIME, AND AVAILABLE DATA SHOULD ALLOW US TO DO SO.

ELEMENT #5: FINALLY, THE FIFTH ELEMENT IS: ONE-ON-ONE EVALUATIVE WORK WITH CONGRESSIONAL COMMITTEE STAFF. THIS WORK FOCUSES ON LEGISLATIVE LANGUAGE MANDATING EVALUATION, ON THE PREPARATION FOR HEARINGS EXAMINING EVALUATIVE QUESTIONS, AND ON AGENCY OVERSIGHT. HERE WE WORK IN-DEPTH WITH A SINGLE COMMITTEE, TRYING TO TRANSFER EVALUATIVE SKILLS, AS WELL AS TO PROVIDE EVALUATION HELP IN A PARTICULAR AREA. THIS WORK WILL BE SUPPLEMENTED BY BRIEFINGS TO INTERESTED CONGRESSIONAL STAFF, TO ACQUAINT THEM WITH ASPECTS OF EVALUATION DESIGN WHICH CAN BE USEFUL TO THEM IN THEIR WORK.

THESE FIVE ELEMENTS, THEN, NOW FURNISH THE ESSENTIAL COMPONENTS OF OUR APPROACH TOWARD TRYING TO INCREASE THE USE OF EVALUATIVE FINDINGS BY THE CONGRESS. THE SHAPE OF THIS APPROACH -- FOCUSED ON THE INFORMATION NEED OF THE USER RATHER THAN ON THE EVALUATIVE INTERESTS OF THE PRODUCER -- HAS EVOLVED FROM AN EFFORT TO AVOID OR ADDRESS THE PROBLEMS ENCOUNTERED BETWEEN LEGISLATIVE USERS AND EVALUATORS IN THE PAST.

THE STRATEGY IS BASED ON THE BELIEF THAT THE ROUTINIZATION OF THESE TWO KINDS OF LINKAGES -- THE DIRECT LINKAGE OF THE EVALUATION TO THE USER NEED, AND CONTINUOUS COMMUNICATION BETWEEN SPONSOR AND EVALUATOR -- SHOULD LEAD TO: MORE RELEVANCE AND TIMELINESS IN THE EVALUATION PRODUCT; NO REDUCTION IN EVALUATIVE QUALITY; A GRADUALLY INCREASING UNDERSTANDING OF THE EVALUATIVE PROCESS BY THE LEGISLATIVE USER; MORE SATISFACTION ON THE PART OF THAT USER; AND AS A RESULT, IMPROVED CONGRESSIONAL USE OF EVALUATIVE FINDINGS.